Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 98 (2007) 1293–1304
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

On the Student's t-distribution and the t-statistic

Zhenhai Yang[a,1], Kai-Tai Fang[b,*], Samuel Kotz[c]

[a] *College of Applied Sciences, Beijing University of Technology, Beijing 10002, China*
[b] *BNU-HKBU United International College, Zhuhai Campus of Beijing Normal University, Jinfenf Road, Zhuhai 519085, China*
[c] *Department of Engineering Management, George Washington University, USA*

Received 26 September 2005
Available online 22 February 2007

Abstract

In this paper we provide rather weak conditions on a distribution which would guarantee that the t-statistic of a random vector of order n follows the t-distribution with $n-1$ degrees of freedom. The results sharpen the earlier conclusions of Mauldon [Characterizing properties of statistical distributions, Quart. J. Math. 2(7) (1956) 155–160] and the more recent advances due to Bondesson [When is the t-statistic t-distributed, Sankhyā, Ser. A 45 (1983) 338–345]. The basic tool involved in the derivations is the vertical density representation originally suggested by Troutt [A theorem on the density of the density ordinate and an alternative interpretation of the Box–Muller method, Statistics 22(3) (1991) 463–466; Vertical density representation and a further remark on the Box–Muller method, Statistics 24 (1993) 81–83]. Several illustrative examples are presented.

AMS 1991 subject classification: 62E10; 62E15

1. Introduction

Let $X_1, \ldots, X_n$ be a random sample from a cumulative distribution function (cdf) $F(\cdot)$ with mean zero and let

$$T_n = T(\mathbf{X}_n) = \frac{\sqrt{n}\bar{X}_n}{s_n}, \tag{1}$$

* Corresponding author. Fax: +852 3411 5811.
E-mail addresses: zhyang@bjpu.edu.cn (Z. Yang), ktfang@uic.edu.hk (K.-T. Fang), kotz@seas.gwu.edu (S. Kotz).
[1] The author is supported by Chinese Natural Science Foundation with contract 10371005 and Beijing Natural Science Foundation.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.11.003

be its t-statistic, where

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2, \quad \mathbf{X}_n = (X_1, \ldots, X_n)^\top.$$

It is well known that if the sample comes from a normal distribution, $\mathcal{N}(0, \sigma^2)$, T_n has the Student t-distribution with $n-1$ degrees of freedom. Many historians of statistics are of the opinion that the most important achievement in statistics in the 20th century was the Pearson's [9] statistic leading to the chi-squared test of fit. In our opinion, the development of the Student's t-statistic by W. S. Gosset ("Student" [10]) was an equally significant contribution that becomes a cornerstone of modern statistical methodology.

Mauldon [8] raised the question for which pdfs the t-statistic, defined by (1), is t-distributed with $n-1$ degrees of freedom. See also an even earlier important contribution of Brown [3]. Along this line the problem at hand is a particular case of the characterization problem of finding all the pdfs for which a certain statistic possesses the property which is a characteristic for these pdfs. Many authors, among them Kagan et al. [7] and Bondesson [1,2], have tackled Mauldon's problem. Bondesson [2] showed that a t-statistic has Student's t-distribution with $n-1$ degrees of freedom for all sample sizes $n \geqslant 2$ if and only if the underlying distribution (assumed to have finite moments of all orders) is normal with mean zero.

It is not necessary that $X_1, \ldots, X_n$ in definition (1) is an independent sample. Indeed consider $X_1, \ldots, X_n$ as a random vector $\mathbf{X}_n = (X_1, \ldots, X_n)^\top$ each component of which having the same marginal distribution function, $F(\cdot)$. Efron [4] has pointed out that the weaker condition of symmetry can replace the normality assumption. Later, Fang et al. [5] have shown that the t-statistic has a t-distribution if the vector $\mathbf{X}_n$ has a spherical distribution. An extension of Mauldon's problem is then to find all $F(\cdot)$ for which the Student's t-statistic has the t-distribution with $n-1$ degrees of freedom.

In some applications of the t-test, $X_1, \ldots, X_n$ may have different marginal distributions. For example, let $\mathbf{Z_n} = (Z_1, \ldots, Z_n)^\top$ be a random sample from $\mathcal{N}(0, 1)$, and let $X_1 = Z_{(1)} \leqslant \cdots \leqslant X_n = Z_{(n)}$ be the order statistics of $Z_1, \ldots, Z_n$. Evidently, $T(\mathbf{X_n}) = T(\mathbf{Z_n})$ and $X_1, \ldots, X_n$ have different marginal distributions and moreover are not independent. This example shows that a more general case could be considered.

Section 2 presents a general theory which allows us to find many interesting cases for which the t-statistic has the t-distribution with $n-1$ degrees of freedom where $X_1, \ldots, X_n$ are either dependent or have different marginal distributions. Section provides many interesting examples. The last section concludes and suggests further study.

Throughout the paper, the norm $\|\cdot\|$ is the L_2-norm, $L_k(A)$ is the Lebesgue measure of A in $\mathbb{R}^k$ and for a $(k-1)$-dimension manifold in $\mathbb{R}^k$, $\bar{L}_k(B)$ denotes the Lebesgue measure of B on this manifold. The uniform distribution on A is denoted by $\mathcal{U}(A)$. The surface of the unit sphere is denoted by $S_n = \{\mathbf{x} : \mathbf{x} \in R^n, \|\mathbf{x}\| = 1\}$, and we shall always assume that $n > 2$.

2. When the t-statistic is t-distributed

In this section, we concern ourselves with the Mauldon's problem in the class of random n-vectors, $\mathbf{X}_n = (X_1, \ldots, X_n)^\top$, designated by $\mathcal{F}$, where $\mathbf{X}_n$ has a continuous density function $f(\mathbf{x}_n)$, $\mathbf{x}_n \in \mathbb{R}^n$, and satisfying the following conditions:

1. $f_0 = f(\mathbf{0}_n) = \sup\{f(\mathbf{x}_n) : \mathbf{x}_n \in \mathbb{R}^n\} < \infty$, $\mathbf{0}_n = (0, \ldots, 0)^\top \in \mathbb{R}^n$;
2. $E(\sum_{k=1}^{n} X_k) = 0$.

In the class $\mathcal{F}$ we are going to derive a sufficient and necessary condition for the t-statistic defined in (1) to have the t-distribution with $n-1$ degrees of freedom.

Set $\mathbf{U}_n = \mathbf{X}_n/\|\mathbf{X}_n\|$. Observe that $\|\mathbf{U}_n\| = 1$ and $T_n(\mathbf{X}_n) = T_n(\mathbf{U}_n)$. We shall proceed as follows: in Section 2.1 the density of $\mathbf{U}_n$ on $\|\mathbf{u}_n\| = 1$ is obtained via the vertical density representation (VDR) theory [11]. In Section 2.2, the density of the t-statistic $T(\mathbf{U}_n)$ is represented in terms of $p(\mathbf{u}_n)$. Section 2.3 provides a sufficient and necessary condition for Mauldon's conjecture to be valid.

2.1. *The density function of* $\mathbf{U}_n$

The VDR theory initially developed by Troutt [11,12] and described in Troutt et al. [13] is used in deriving the density of $\mathbf{U}_n$. In particular, we need to recall some results on VDR in Fang et al. [6]. Some applications of VDR can refer to Yang et al. [14]. Let $f(\mathbf{x}_n)$, $\mathbf{x}_n \in \mathbb{R}^n$ be a density function defined on $\mathbb{R}^n$. Set

$$\begin{aligned} D_{n,[f]} &= \{\mathbf{x}_{n+1} = (\mathbf{x}_n^\top, x_{n+1})^\top : 0 < x_{n+1} \leqslant f(\mathbf{x}_n)\}, \\ D_{n,[f]}(v) &= \{\mathbf{x}_n : f(\mathbf{x}_n) \geqslant v\}, \quad 0 < v \leqslant f_0 = \sup\{f(\mathbf{x}_n), \mathbf{x}_n \in \mathbb{R}^n\}. \end{aligned} \tag{2}$$

If $f \in \mathcal{F}$, then we have

$$\mathbf{0}_n \in D_{n,[f]}(v) \quad \forall v \in (0, f_0].$$

It can be shown that the Lebesgue measure of $D_{n,[f]}$ equals

$$1 = L_{n+1}(D_{n,[f]}) = \int_0^{f_0} L_n(D_{n,[f]}(v))\, dv.$$

The following lemma is taken from Fang et al. [6].

Lemma 2.1. *Let* $(X_1, \ldots, X_n, X_{n+1}) = (\mathbf{X}_n^\top, X_{n+1})$ *be a random* $(n+1)$*-vector that is uniformly distributed on* $D_{n,[f]}$*. Then we have*

(1) *The density function of* X_{n+1} *is*

$$L_n(D_{n,[f]}(v)), \quad 0 < v \leqslant f_0 = \sup\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\} \text{ and zero otherwise},$$

where $D_{n,[f]}(v)$ *is defined in* (2);

(2) *The density function of* $\mathbf{X}_n$ *is* $f(\cdot)$;

(3) *Given* $X_{n+1} = v$, *the conditional distribution of* $\mathbf{X}_n$ *is the uniform distribution on* $D_{n,[f]}(v)$.

Furthermore, (1) *and* (3) *imply* (2).

For any random variable $R > 0$ with a density function $f_R(r)$, we have

$$P(\mathbf{X}_n \leqslant \mathbf{x}) = \int_0^\infty P(\mathbf{X}_n \leqslant \mathbf{x} | R = r) f_R(r)\, dr, \quad \mathbf{x} \in R^n. \tag{3}$$

In particular, we choose

$$R = R^* = \sqrt{-2\ \ln(X_{n+1}/f_0)}, \quad 0 < X_{n+1} \leqslant f_0, \tag{4}$$

and it follows that $X_{n+1} = f_0 \exp\{-\frac{1}{2}(R^*)^2\}$. We observe that the density function of R^* can easily be obtained by means of Lemma 2.1 as follows:

$$f_{R^*}(r) = f_0 L_n(D_{n,[f]}(f_0 \exp(-\tfrac{1}{2}r^2))) \cdot r e^{-1/2r^2}, \quad 0 \leqslant r < \infty.$$

Utilizing (3) the density function of $\mathbf{X}_n$ can be represented as

$$f(\mathbf{x}) = \int_{\{r:\mathbf{x}\in D_{n,[f]}(f_0\ \exp(-\frac{1}{2}r^2))\}} \frac{f_{R^*}(r)}{L_n(D_{n,[f]}(f_0\ \exp(-\frac{1}{2}r^2)))}\, dr$$

$$= f_0 \int_{\{r:\mathbf{x}\in D_{n,[f]}(f_0\ \exp(-\frac{1}{2}r^2))\}} re^{-1/2r^2}\, dr.$$

Note that from (4)

$$\min\left\{r : \mathbf{x} \in D_{n,[f]}\left(f_0\ \exp\left(-\frac{1}{2}r^2\right)\right)\right\} = \sqrt{-2\ \ln \frac{f(\mathbf{x})}{f_0}}, \quad \text{and}$$

$$\text{the range of } \left\{r : \mathbf{x} \in D_{n,[f]}\left(f_0\ \exp\left(-\frac{1}{2}r^2\right)\right)\right\} \text{ is } \left(\sqrt{-2\ln \frac{f(\mathbf{x})}{f_0}}, \infty\right).$$

Finally, we have an expression for $f(\mathbf{x})$

$$f(\mathbf{x}) = f_0 \int_{\{r:\mathbf{x}\in D_{n,[f]}(f_0\ \exp(-\frac{1}{2}r^2))\}} re^{-1/2r^2}\, dr = f_0 \int_{\sqrt{-2\ \ln \frac{f(\mathbf{x})}{f_0}}}^{\infty} re^{-1/2r^2}\, dr.$$

Summarizing the above discussion we have the following lemma where the class $\mathcal{F}$ is defined at the very beginning of this section.

Lemma 2.2. *Let $\mathbf{X}_n$ be a random vector in $\mathcal{F}$. Its cdf can be represented as* (3), *where*

(1) *R is a non-negative random variable with the density function*

$$g(r) = L_n(D_{n,[f]}(f_0\ \exp(-\tfrac{1}{2}r^2)))re^{-1/2r^2}, \quad 0 \leqslant r < \infty;$$

(2) *Given $R = r$, the conditional distribution of $\mathbf{X}_n$ is the uniform distribution $U(D^*_{n,[f]}(r))$, where*

$$D^*_{n,[f]}(r) = D_{n,[f]}(f_0\ \exp(-\tfrac{1}{2}r^2)).$$

The derivation of the density of $\mathbf{U}_n = \mathbf{X}_n/\|\mathbf{X}_n\|$ is given in the next lemma.

Lemma 2.3. *Let $\mathbf{X}_n$ be a random vector in $\mathcal{F}$ and $\mathbf{U}_n = \mathbf{X}_n/\|\mathbf{X}_n\|$. Then the density function of $\mathbf{U}_n$ on S_n, the surface of the unit sphere in R^n, is given by*

$$p(\mathbf{u}_n) = \int_0^{\infty} q(\mathbf{u}_n; r)re^{-1/2r^2}\, dr, \tag{5}$$

where

$$q(\mathbf{u}_n; r) = \int_{A(\mathbf{u}_n;r)} s^{n-1}\, ds, \tag{6}$$

and

$$A(\mathbf{u}_n; r) = \{s : s\mathbf{u}_n \in D^*_{n,[f]}(r)\} \quad \forall \mathbf{u}_n \in S_n.$$

Proof. Let R be random variable with density function

$$g(r) = L_n(D_{n,[f]}(f_0\ \exp(-\tfrac{1}{2}r^2)))re^{-1/2r^2}, \quad 0 \leqslant r < \infty.$$

By Lemma 2.2, for a given $R = r$, the conditional distribution of $\mathbf{X}_n$ is the uniform distribution on $D^*_{n.[f]}(r)$. Set the neighborhood of $\mathbf{u}_n$ to be

$$\Delta(\mathbf{u}_n, \delta) = \{\mathbf{x}_n : \|\mathbf{x}_n - \mathbf{u}_n\| \leqslant \delta, \|\mathbf{x}_n\| = 1\} \quad \forall \|\mathbf{u}_n\| = 1, \quad \delta > 0, \tag{7}$$

then the following steps are a direct consequence of (5)–(7)

$$\begin{aligned}
p(\mathbf{u}_n|R = r) &= \lim_{\delta \longrightarrow 0} \frac{P(\mathbf{U}_n \in \Delta(\mathbf{u}_n, \delta)|R = r)}{\bar{L}_{n-1}(\Delta(\mathbf{u}_n, \delta))} \\
&= \lim_{\delta \longrightarrow 0} \frac{P(\mathbf{X}_n \in \{\mathbf{y}_n : \mathbf{y}_n = s\mathbf{x}_n \in D^*_{n,[f]}(r), s > 0,\ \mathbf{x}_n \in \Delta(\mathbf{u}_n, \delta)\}|R = r)}{\bar{L}_{n-1}(\Delta(\mathbf{u}_n, \delta))} \\
&= \lim_{\delta \longrightarrow 0} \frac{L_n(\{\mathbf{y}_n : \mathbf{y}_n = s\mathbf{x}_n \in D^*_{n,[f]}(r), s > 0,\ \mathbf{x}_n \in \Delta(\mathbf{u}_n, \delta)\})}{L_n(D^*_{n,[f]}(r))\bar{L}_{n-1}(\Delta(\mathbf{u}_n, \delta))} \\
&= \frac{1}{L_n(D^*_{n,[f]}(r))} \\
&\quad \times \lim_{\delta \longrightarrow 0} \int_{A(\mathbf{u}_n;r)} \frac{\bar{L}_{n-1}(\{\mathbf{u}_n : \mathbf{u}_n \in \Delta(\mathbf{u}_n, \delta), s\mathbf{u}_n \in D_{n,[f]}(r)\})}{\bar{L}_{n-1}(\Delta(\mathbf{u}_n, \delta))} s^{n-1}\, ds \\
&= \frac{1}{L_n(D^*_{n,[f]}(r))} \int_{A(\mathbf{u}_n;r)} s^{n-1}\, ds.
\end{aligned}$$

Hence

$$p(\mathbf{u}_n) = \int_0^\infty p(\mathbf{u}_n|R = r) g(r)\, dr = \int_0^\infty q(\mathbf{u}_n; r) r e^{-\frac{1}{2}r^2}\, dr$$

as claimed. □

Example 2.1. When $\mathbf{X}_n = R\mathbf{V}_n$ where $R > 0$, $\mathbf{V}_n$ is uniformly distributed on a compact set $D \subset R^n$, then the density function of $\mathbf{U}_n = \mathbf{X}_n/\|\mathbf{X}_n\| = \mathbf{V}_n/\|\mathbf{V}_n\|$ is

$$p(\mathbf{u}_n) = \frac{1}{L_n(D)} \int_{A(\mathbf{u}_n;\, D)} s^{n-1}\, ds, \quad \|\mathbf{u}_n\| = 1, \tag{8}$$

where $A(\mathbf{u}_n; D) = \{s : s > 0, s\mathbf{u}_n \in D, \|\mathbf{u}_n\| = 1\}$. Obviously the density in (5) is a special case of (8).

2.2. The density function of the t-statistic defined in (1)

Now we derive the density function of T_n via the density function of $\mathbf{U}_n$. Denote

$$e(n, t) = \frac{\frac{t}{\sqrt{n-1}}}{\sqrt{1 + \frac{t^2}{(n-1)}}}.$$

We then have

$$\begin{aligned}
P(T_n(\mathbf{X}_n) \leqslant t) &= P(T_n(\mathbf{U}_n) \leqslant t) = EI_{(-\infty,t]}(T_n(\mathbf{U}_n)) \\
&= EI_{[-1,e(n,t)]}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n U_i\right).
\end{aligned} \tag{9}$$

Let $\mathbf{Q}$ be an orthogonal matrix of order n whose last row $(1/\sqrt{n}, \dots, 1/\sqrt{n})$ and set

$$\mathbf{V}_n = \mathbf{Q}\mathbf{U}_n.$$

Obviously $\|\mathbf{V}_n\| = 1$. The density function $p^*(\mathbf{v}_n)$ of $\mathbf{V}_n$ is

$$p^*(\mathbf{v}_n) = p(Q^\top \mathbf{v}_n), \quad \mathbf{v}_n \in S_n,$$

and the cdf of T_n is then

$$\begin{aligned}
F_{T_n}(t) &= P(T_n(\mathbf{X}_n) \leqslant t) = EI_{[-1,e(n,t)]}(V_n) \\
&= \int_{\{\sum_{i=1}^n v_i^2=1\}\cap\{\mathbf{v}_n:-1\leqslant v_n\leqslant e(n,t)\}} p^*(\mathbf{v}_n)\bar{L}_n(d\mathbf{v}_n) \\
&= \int_{-1}^{e(n,t)} (1-v_n^2)^{-\frac{1}{2}}\, dv_n \int_{\{\sum_{i=1}^{n-1} v_i^2=1-v_n^2\}} p^*(v_1,\dots,v_{n-1},v_n)\bar{L}_{n-1}(d\mathbf{v}_{n-1}) \\
&= \int_{-1}^{e(n,t)} (1-v_n^2)^{\frac{n-3}{2}}\, dv_n \int_{\{\sum_{i=1}^{n-1} v_i^2=1\}} p^*\left(\sqrt{1-v_n^2}\mathbf{v}_{n-1}, v_n\right)\bar{L}_{n-1}(d\mathbf{v}_{n-1}) \\
&= \int_{-1}^{e(n,t)} (1-v_n^2)^{\frac{n-3}{2}} h_n(v_n)\, dv_n, \qquad (10)
\end{aligned}$$

where

$$h_n(v) = \int_{\{\|\mathbf{v}_{n-1}\|=1\}} p^*\left(\sqrt{1-v_n^2}\mathbf{v}_{n-1}, v\right)\bar{L}_{n-1}(d\mathbf{v}_{n-1}). \qquad (11)$$

From (10), we have the following lemma:

Lemma 2.4. *The probability density of the random variable*

$$V_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n U_i = \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i}{\|\mathbf{X}_d\|}$$

is

$$p_e(v) = (1-v^2)^{\frac{n-3}{2}} h_n(v), \quad -1\leqslant v\leqslant 1.$$

Example 2.2. Let us verify the correctness of (10) and (11) by using a well-known case. When $\mathbf{X}_n$ has independent components each having a $\mathcal{N}(0,\sigma^2)$ distribution, and thus $T(\mathbf{X}_n)$ follows the t-distribution with $n-1$ degrees of freedom. In this case, $D^*_{n,[f]}(r) = \{\mathbf{x}_n : \|\mathbf{x}_n\|\leqslant r\}$, both $\mathbf{V}_n$ and $\mathbf{U}_n$ are uniformly distributed, on the unit sphere and on the surface of the unit sphere, respectively, and we obtain from (11) and (10) that

$$h_n(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{n/2}}\frac{2\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)},$$

$$F_{T_n}(t) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}\int_{-1}^{e(n,t)} (1-v^2)^{\frac{n-3}{2}}\, dv. \qquad (12)$$

Hence the density function associated with $F_{T_n}(t)$ is

$$g_{n-1}(t) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}\left(1-\frac{\frac{t^2}{n-1}}{\sqrt{1+\frac{t^2}{(n-1)}}}\right)^{\frac{n-3}{2}}\left(\frac{\frac{1}{\sqrt{n-1}}}{\sqrt{1+\frac{t^2}{(n-1)}}}-\frac{\frac{t^2}{[\sqrt{n-1}]^3}}{\left[\sqrt{1+\frac{t^2}{(n-1)}}\right]^3}\right)$$
$$= \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{(n-1)\pi}\Gamma\left(\frac{n-1}{2}\right)}\left(1+\frac{t^2}{n-1}\right)^{\frac{-n}{2}}$$

which is just the t-distribution with $n-1$ degrees of freedom.

2.3. A sufficient and necessary condition for a t-statistic to be t-distributed

Theorem 2.1. *Let $\mathbf{X}_n, n \geqslant 3$ be a random vector in $\mathcal{F}$. The t-statistic $T(\mathbf{X}_n)$ has the Student's t-distribution with $n-1$ degrees of freedom if and only if*

$$h_n(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}, \quad -1 < v < 1, \tag{13}$$

where $h_n(v)$ is defined by (11).

Proof. It is easy to obtain from (10) and (12) that a necessary and sufficient condition of the t-statistic to have the t-distribution with $(n-1)$ degrees of freedom is

$$\int_{-1}^{e(n,t)} (1-v^2)^{\frac{n-3}{2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}\, dv = \int_{-1}^{e(n,t)} (1-v^2)^{\frac{n-3}{2}} h_n(v)\, dv \quad \forall t \in R.$$

That implies

$$h_n(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}, \quad -1 < v < 1,$$

as claimed. □

By Lemma 2.4, Theorem 2.1 can be equivalently stated as follows.

Theorem 2.1′. *Let $\mathbf{X}_n, n \geqslant 3$ be a random vector in $\mathcal{F}$. The t-statistic $T(\mathbf{X}_n)$ has the Student's t-distribution with $n-1$ degrees of freedom if and only if the density function of $(1/\sqrt{n})\sum_{i=1}^n X_i/\|\mathbf{X}_n\|$ is*

$$p_e(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}(1-v^2)^{\frac{n-3}{2}}, \quad -1 \leqslant v \leqslant 1. \tag{14}$$

3. Some examples

In this section, some interesting examples that have some potential applications are presented.

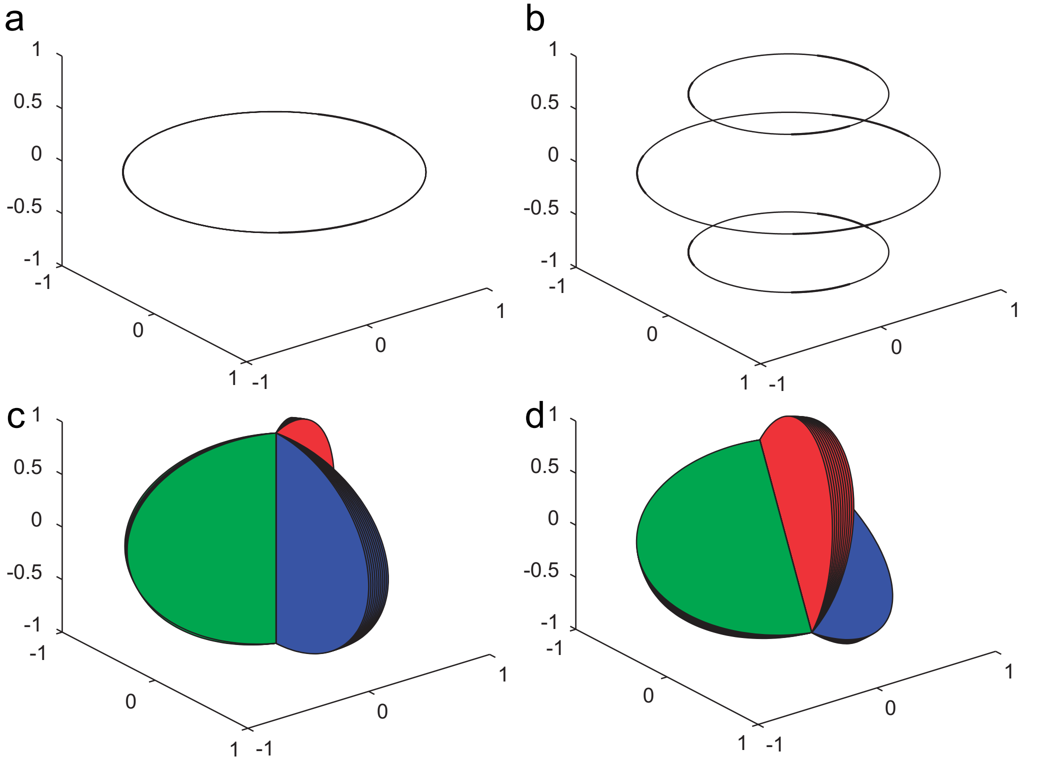

Fig. 1. (a) $\bar{D}_{n-1}$, (b) $\bar{D}_n$, (c) D_n and (d) $D_{n,C}$.

Example 3.1. Let $\bar{S}_{n-1}$ be the surface of a unit sphere in $\mathbb{R}^{n-1}$ to be presented in $\mathbb{R}^n$

$$\bar{S}_{n-1} = \left\{ \mathbf{u}_n(0) = (u_1, \ldots, u_{n-1}, 0)^\top : \sum_{i=1}^{n-1} u_i^2 = 1 \right\}$$

and let $\bar{D}_{n-1}$ be a subset of $\bar{S}_{n-1}$ with positive Lebesgue measure, $\bar{L}_{n-1}(\bar{D}_{n-1}) > 0$. Let $\mathbf{e}_n = (0, \ldots, 0, 1)^\top \in \mathbb{R}^n$ and set

$$\bar{D}_n = \left\{ \mathbf{u}_n = \sqrt{1 - u_n^2}\, \mathbf{u}_n(0) + u_n \mathbf{e}_n : \mathbf{u}_n(0) \in \bar{D}_{n-1}, -1 \leqslant u_n \leqslant 1 \right\},$$

$$D_n = \{ \mathbf{v}_n = s\mathbf{u}_n : 0 \leqslant s \leqslant 1, \mathbf{u}_n \in \bar{D}_n \},$$

$$D_{n,C} = C^\top D_n = \{ \mathbf{x}_n = C^\top \mathbf{u}_n : \mathbf{u}_n \in D_n \},$$

where C is an orthogonal matrix with the last row $(1/\sqrt{n}, \ldots, 1/\sqrt{n})$. Obviously, $\bar{D}_n$ is a subset of $\bar{S}_n$, the surface of the unit sphere in R^n, and D_n is the set that includes all points $s\mathbf{u}_n$, $0 \leqslant s \leqslant 1$. The sets $\bar{D}_{n-1}$ and $\bar{D}_n$ are called a *base surface set* and a *generating surface set*, respectively. Obviously, $D_n = CD_{n,C}$ and $CD_{n,C}$ is called a *base transformed space*. Fig. 1 shows $\bar{D}_{n-1}$, $\bar{D}_n$, D_n and $D_{n,C}$, where (a) stands for $\bar{D}_{n-1}$; the corresponding $\bar{D}_n$ for the $\mathbf{u}_n$-values are presented in (b); its D_n and $D_{n,C}$ are showed in (c) and (d), respectively. The shapes of D_n and $D_{n,C}$ are like three pieces of a watermelon.

Let $\mathbf{X}_n = R\mathbf{V}_n$, where $R \geqslant 0$ and $\mathbf{V_n}$ are independent, and $\mathbf{V_n}$ is uniformly distributed on the base sample space $D_{n,C}$. Thus, the transform $\mathbf{V}_n^* = C\mathbf{V}_n$ is uniformly distributed on the base sample space D_n with $E(V_n^*) = 0$, where V_n^* is the last component of $\mathbf{V}_n$. Note $\mathbf{e}_n = (0, \dots, 0, 1)^\top = C\mathbf{1}_n$, where $\mathbf{1}_n = (1/\sqrt{n}, \dots, 1/\sqrt{n})^\top$. We thus have

$$E\left\{\frac{1}{\sqrt{n}}\sum_{I=1}^{n} X_i\right\} = E\{\mathbf{1}^\top \mathbf{X}_n\} = ER \cdot E\{\mathbf{1}^\top \mathbf{V}_n\} = ER \cdot E\{[C\mathbf{1}]^\top C\mathbf{V}_n\}$$

$$= ER \cdot E\{\mathbf{e}^\top \mathbf{V}_n^*\} = ER \cdot EV_n^* = 0.$$

By Lemma 2.3, $\mathbf{U}_n = \mathbf{X}_n/\|\mathbf{X}_n\| = \mathbf{V}_n/\|\mathbf{V}_n\|$ is uniformly distributed on $D_{n,C}$ because

$$p(\mathbf{u}_n) = \frac{1}{L_n(D_{n,c})}\int_0^1 s^{n-1}\,ds = \frac{1}{nL_n(D_n)} = \frac{1}{\bar{L}_n(\bar{D}_n)}, \quad \mathbf{u}_n \in \bar{D}_n;$$

and $p(\mathbf{u}_n) = 0$ elsewhere. From (11), we have

$$h_n(u) = \frac{\int_{\{\|\mathbf{v}_{n-1}\|\}} \bar{L}_{n-1}(d\mathbf{v}_{n-1})}{\bar{L}_n(\bar{D}_n)} = \frac{\bar{L}_n(\bar{D}_{n-1})}{\bar{L}_n(\bar{D}_n)} = \frac{\bar{L}_n(\bar{S}_{n-1}^2)}{\bar{L}_n(\bar{S}_n^2)} = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)},$$

and by Theorem 2.1 the t-statistic $T(\mathbf{X}_n)$ is t-distributed with $n-1$ degrees of freedom.

Note that the random vector $\mathbf{X}_n$ can be also represented as

$$\mathbf{X}_n = R\mathbf{V}_n = R\|\mathbf{V}_n\|\frac{\mathbf{V}_n}{\|\mathbf{V}_n\|} = R^*\mathbf{U}_n$$

and the derivation works equivalently for both $\mathbf{U}_n$ and $\mathbf{V}_n$.

Remark. The process of generating $D_{n,C}$ from D_{n-1} is as follows:

Step 1: Choose

$$\bar{D}_{n-1} \subseteq \bar{S}_{n-1} = \left\{\mathbf{u}_n(0) = (u_1, \dots, u_{n-1}, 0)^\top : \sum_{k=1}^{n-1} u_k^2 = 1\right\}$$
$$= \{\mathbf{u}_n : \mathbf{e}_n^\top \mathbf{u}_n = 0,\ \|\mathbf{u}_n\| = 1\}.$$

Step 2: Generate $\bar{D}_n$ by reducing the radius and parallel movement

$$\bar{D}_n = \left\{\sqrt{1-v^2}\mathbf{u}_n(0) + v\mathbf{e}_n : \mathbf{u}_n(0) \in \bar{D}_{n-1},\ -1 \leqslant v \leqslant 1\right\}.$$

Step 3: Generate D_n by adding segments: $D_n = \{\mathbf{v}_n = s\mathbf{u}_n :, 0 \leqslant s \leqslant 1, \mathbf{u}_n \in \bar{D}_n\}$.

Step 4: Generate $D_{n,C}$ by the orthogonal transformation $D_{n,C} = C^\top D_n$.

In fact, the orthogonal transformation at the last step can be directly applied to $\bar{D}_{n-1}$:

$$\bar{D}_{n-1,C} = C^\top \bar{D}_{n-1} = \{C^\top \mathbf{u}_n : \mathbf{1}^\top \mathbf{u}_n = 0, \|\mathbf{u}_n\| = 1, \mathbf{u}_n \in \bar{D}_{n-1}\}.$$

It is easy to see that

$$\bar{D}_{n,C} = C^\top \bar{D}_n$$

and the set $D_{n,C}$ can be constructed directly by moving a subset of $D_{n,C}$ along the direction $\mathbf{1}_n$. Also the process of generating $D_{n,C}$ from $\bar{D}_{n-1,C}$ is as follows:

Step 1: Choose $\bar{D}_{n-1,C} \subseteq \{\mathbf{u}_n : \mathbf{1}_n^\top \mathbf{u}_n = 0,\ \|\mathbf{u}_n\| = 1\}$.

Step 2: Generate $\bar{D}_{n,C}$ by reducing the radius and parallel movement

$$\bar{D}_{n,C} = \left\{\mathbf{y}_n = \left(\sqrt{1-v^2}\mathbf{u}_n + v\mathbf{1}_n\right),\ \mathbf{u}_n \in \bar{D}_{n-1,C},\ -1 \leqslant v \leqslant 1\right\}.$$

Step 3: Generate $\bar{D}_{n,C}$ by adding segments:

$$D_{n,C} = \{\mathbf{x}_n = s\mathbf{u}_n : 0 \leqslant s \leqslant 1, \mathbf{u}_n \in \bar{D}_{n,C}\}.$$

The two processes of constructing set $D_{n,C}$ as described above are equivalent. We now present several cases of $\bar{D}_{n-1}$.

Case 1: *Angle base set*. Let both $\mathbf{a}_{n-1}$ and $\mathbf{b}_{n-1}$ be unit column vectors in $R^{(n-1)} = \{(x_1, \ldots, x_{n-1}, 0)\}$ and define two hyperplanes:

$$l_1 : \mathbf{a}_{n-1}^\top \mathbf{x}_{n-1} = 0, \quad \mathbf{x}_{n-1} \in \mathbb{R}^{(n-1)},$$
$$l_2 : \mathbf{b}_{n-1}^\top \mathbf{x}_{n-1} = 0, \quad \mathbf{x}_{n-1} \in \mathbb{R}^{(n-1)}.$$

Choose $\mathbf{a}_{n-1}$ and $\mathbf{b}_{n-1}$ such that the set

$$\bar{D}_{n-1} = \{\mathbf{x}_{n-1} : \|\mathbf{x}_{n-1}\| = 1, \mathbf{a}_{n-1}^\top \mathbf{x}_{n-1} \leqslant 0 \leqslant \mathbf{b}_{n-1}^\top \mathbf{x}_{n-1}, \mathbf{x}_{n-1} \in \mathbb{R}^{(n-1)}\}$$

is not empty. We can construct also the set corresponding to the set $D_{n,C}$ and a random vector $\mathbf{X}_n = R\mathbf{V}_n$ as described above. Here, $D_{n,C}$ is like a piece of a watermelon. Obviously, its t-statistic $T(\mathbf{X}_n)$ obeys the t-distribution with $n-1$ degrees of freedom.

Case 2: *Order statistics of a normal sample*. Let $\mathbf{X}_n = (X_1, \ldots, X_n)^\top$ be a sample from a normal distribution $N(0, 1)$ and $\mathbf{X}_{(n)} = (X_{(1)}, \ldots, X_{(n)})^\top$ be its order statistic. Evidently we have $T(\mathbf{X}_n) = T(\mathbf{X}_{(n)})$ and both statistics obey the t-distribution with $n-1$ degrees of freedom. However, the distributions of $\mathbf{X}_n$ and $\mathbf{X}_{(n)}$ are quite different. How to explain this phenomenon?

In the following we show that the t-statistic generated by the above order statistics can be produced from some chosen $\bar{D}_n$. From the representation of $\mathbf{X}_n = R\mathbf{U}_n$, we have

$$\mathbf{X}_{(n)} = R\mathbf{U}_{(n)}, \quad T(\mathbf{X}_{(n)}) = T(\mathbf{U}_{(n)}),$$

where $\mathbf{U}_{(n)}$ is the order statistics of $\mathbf{U}_n$. The latter has the uniform distribution on S_n. Set

$$\bar{D}_{n-1} = \{\mathbf{x}_n = (x_1, \ldots, x_n) : \|\mathbf{x}_n\| = 1, \mathbf{1}_n^\top \mathbf{x}_n = 0,\ x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n\}.$$

Construct $\bar{D}_{n-1}(v)$ by the set $\bar{D}_{n-1}$ along the direction $\mathbf{1}_n$:

$$\bar{D}_{n-1}(v) = \{\mathbf{x}_n = (x_1, \ldots, x_n) : \|\mathbf{x}_n - v\mathbf{1}_n\| = 1, \mathbf{1}_n^\top \mathbf{x}_n = \sqrt{n}\, v, x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n\}.$$

The idea for construction of $\bar{D}_{n-1}(v)$ is by a one-to-one mapping $M(\cdot)$ between the two sets

$$M(\mathbf{x}_n) = \mathbf{x}_n + v\mathbf{1}_n = \mathbf{y}_n \in \bar{D}_{n-1}(v) \quad \forall \mathbf{x}_n \in \bar{D}_{n-1}.$$

Actually this mapping is $\{\mathbf{1}_n^\top \mathbf{x}_n = 0\} \Leftrightarrow \{\mathbf{1}_n^\top \mathbf{x}_n = \sqrt{n}v\}$, $v \in R$. A useful property of the mapping is that orders of the components are preserved. Define

$$\bar{D}_n = \left\{\mathbf{y}_n = \left(\sqrt{1-v^2}\mathbf{x}_n\right) + v\mathbf{1}_n : -1 \leqslant v \leqslant 1, \mathbf{x}_n \in D_{n-1}\right\}.$$

We can see that the set of the order statistics of $\mathbf{U}_n$, $\{\mathbf{x}_n : \|\mathbf{x}_n\| = 1, \quad x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n\}$, is the generating surface set of $\bar{D}_{n-1}$. Hence the case of order statistics is a particular case of Example 3.1.

Case 3: *Cone base set.* Set

$$\bar{D}_{n-1}(d) = \{\mathbf{v}_{n-1} : \|\mathbf{v}_{n-1}\| = 1, \|\mathbf{v}_{n-1} - \mathbf{c}_{n-1}\| \leqslant d\},$$

where

$$\mathbf{c}_{n-1} \in R^{n-1} \text{ is a constant vector,} \quad \|\mathbf{c}_{n-1}\| = 1, 0 < d \leqslant \sqrt{2}.$$

Using the above procedure, we can define a new random vector $\mathbf{X}_n$ such that the corresponding $T(\mathbf{X}_n)$ obeys the t-distribution with $n-1$ degrees of freedom.

Example 3.2. In the previous example the generating surface set was constructed solely by parallel movement. Now we shall construct generating base surface by means of parallel movement and rotation. Set

$$S_2^* = \{(x, y, 0)^\top : x^2 + y^2 = 1, x, y \geqslant 0, x \leqslant ay\}, \quad a \text{ is a constant, and}$$
$$S_3^* = \left\{(x, y, z)^\top : -1 \leqslant z \leqslant 1, \sqrt{1-z^2} C_{[3]}(z)(x, y, 0)^\top + (0, 0, z)^\top, (x, y, 0)^\top \in S_2^*\right\},$$

where the rotation transforms are determined by

$$C_{[3]}(z) = \begin{bmatrix} C_{[2]}(z) & \begin{matrix} 0 \\ 0 \end{matrix} \\ 0 \; 0 & 1 \end{bmatrix}, \quad C_{[2]}(z) = \begin{bmatrix} \cos \dfrac{k\pi z}{2} & \sin \dfrac{k\pi z}{2} \\ -\sin \dfrac{k\pi z}{2} & \cos \dfrac{k\pi z}{2} \end{bmatrix},$$

and k is a parameter that determines the rotation angle of the base set. Now set

$$\bar{S}_3 = \{(x, y, z)^\top = s(u, v, w)^\top : 0 \leqslant s \leqslant 1, (u, v, w)^\top \in S_3^*\}.$$

Let C is a 3×3 orthogonal matrix whose last row $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ and $\mathbf{V}_3$ be a random vector with a uniform distribution on $S_{3,C} = C^\top \bar{S}_3$. Set $\mathbf{X}_3 = R\mathbf{V}_3$, where R, as above, is a nonnegative random variable and is independent of $\mathbf{V}_3$. Then $\mathbf{U}_3 = \mathbf{V}_3/\|\mathbf{V}_3\|$ is uniformly distributed on $C^\top \bar{S}_3$.

We shall verify that in this case

$$h_3(v) = \frac{\Gamma\left(\frac{3}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{2}{2}\right)} = \frac{1}{2}, \quad -1 < v < 1.$$

Indeed

$$L_3(S_3) = \int_{-1}^{1} \left[\frac{1}{1-z^2}\right]^{\frac{3-3}{2}} dz \int_{\{C_2(z)S_2^*\}} \bar{L}_2(d(x, y)) = 2\int_{S_2^*} \bar{L}_2(d(x, y)) = 2\bar{L}_2(S_2^*),$$

where the last is on the manifold. Also

$$\int_{\{x^2+y^2=1\}} p^*\left(\sqrt{1-z^2(x, y), z}\right) \bar{L}_2(d(x, y)) = \int_{\bar{S}_2} \bar{L}_2(d(x, y)) = \bar{L}_2(\bar{S}_2).$$

From (11) and the definition of $h_n(v)$, the ratio of above two values is $h_3(v) = \frac{1}{2}$, hence it follows that the t-statistic $T(\mathbf{X}_n)$ possesses the t-distribution with 2 degrees of freedom.

4. Discussion

We have obtained the sufficient and necessary condition for t-statistic in the class of $\mathcal{F}$ to be t-distributed. Several non-traditional examples are given. For further study the inverse problem is the most interesting one.

Let $\mathbf{X}_n = (X_1, \ldots, X_n)^\top$ be a random sample (i.e. $X_1, \ldots, X_n$ are i.i.d.) from the population with a probability density $f(\cdot)$. Under the assumption that all the moments exist, Bondesson [2] in his seminal paper proved that if the t-statistic $T(\mathbf{X}_n)$ of the sample follows the t-distribution with $n-1$ degrees of freedom, then the sample comes from a normal distribution. This is called the inverse problem for the t-distribution. Bondesson required the existence of moments of any order of the population distribution. Can we give an answer to the inverse problem under rather weak condition required by Bondesson? This is an open question.

Acknowledgment

The authors would like to thank valuable comments from the Editor and two referees. We also express our gratitude to Prof. Chen, Xuru for his important and useful comments.

References

[1] L. Bondesson, Characterizations of probability laws through constant regression, Z. Wahrscheinlichkeitstheorie verw. Gebiete 30 (1974) 93–115.
[2] L. Bondesson, When is the t-statistic t-distributed, Sankhyā, Ser. A 45 (1983) 338–345.
[3] G.W. Brown, Reduction of a certain class of statistical hypotheses, Ann. Math. Statist. 11 (1940) 254–270.
[4] B. Efron, Student's t-test under symmetry conditions, J. Amer. Statist. Assoc. 64 (1969) 1278–1302.
[5] K.T. Fang, K.W. Ng, S. Kotz, Symmetric Multivariate and Related Distributions, Chapman & Hall, London, 1990.
[6] K.T. Fang, Z.H. Yang, S. Kotz, Generation of multivariate distributions by vertical density representation, Statistics 35 (2001) 281–293.
[7] A.M. Kagan, Y.V. Linnik, C.R. Rao, Characterization Problems in Mathematical Statistics, Wiley, New York, 1973.
[8] J.G. Mauldon, Characterizing properties of statistical distributions, Quart. J. Math. 2 (7) (1956) 155–160.
[9] K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Philos. Mag. (5th Series) 50 (1900) 339–357 Correction, Philos. Mag. (6th Series), 1, 670–671.
[10] "Student" On the probable error of the mean, Biometrika 6 (1908) 1–25.
[11] M.D. Troutt, A theorem on the density of the density ordinate and an alternative interpretation of the Box–Muller method, Statistics 22 (3) (1991) 463–466.
[12] M.D. Troutt, Vertical density representation and a further remark on the Box–Muller method, Statistics 24 (1993) 81–83.
[13] M.D. Troutt, W.K. Pang, S.H. Hou, Vertical Density Representation and its Applications, World Scientific Publishing, Singapore, 2004.
[14] Z.H. Yang, S. Kotz, Center-similar distributions with applications in multivariate analysis, Statist. Probab. Lett. 64 (2003) 335–345.